Pl@ntNet Mobile 2014: Android port and new features

Hervé Goëau
INRIA, ZENITH

Pierre Bonnet
CIRAD, UMR AMAP

Alexis Joly
INRIA, ZENITH

Antoine Affouard
CIRAD, UMR AMAP

Vera Bakić
INRIA, ZENITH

Julien Barbe
INRA, UMR AMAP

Sammuel Dufour
INRA, UMR AMAP

Souheil Selmi
INRIA, ZENITH

Itheri Yahiaoui
Reims university, CReSTIC

Christel Vignau
Tela Botanica

Daniel Barthelemy
CIRAD, BIOS

Nozha Boujemaa
INRIA, Saclay

ABSTRACT

This paper presents consistent updates of Pl@ntNet[1], an image sharing and retrieval application for the identification of plants. New features over the previous iPhone version described in [7] include (i) its porting to most android platforms (ii) more than three times more data (iii) the use of metadata additionally to the visual content in the identification process (iv) a new multi-organ, multi-image and multi-features fusion strategy using separated indexes for each visual feature (v) the integration of cross-languages functionalities. Additionally, the paper presents the new results achieved by our system within ImageCLEF plant identification task 2013 as well as real-world user trials and feedbacks.

1. PL@NTNET VISION AND APPROACH

Building accurate knowledge of the identity, geographic distribution and uses of plants is essential if agricultural development is to be successful and biodiversity is to be conserved. One big challenge, expressed as the *taxonomic gap*, is that identifying plant species is usually impossible for the general public, and often a difficult task for professionals, such as farmers (who have to fight against weed species) or foresters. In this context, content-based visual identification tools are considered as one of the most promising solution [8, 4, 19, 11] particularly mobile applications [16, 1] that could help setting-up massive ecological monitoring systems. A noticeable progress in this way has been achieved by the US consortium at the origin of LeafSnap[2] through the development of an iPhone application allowing a fair identification of 184 common american species based on uncluttered images of leaves and shape boundary features [16].

But this is still far from a real-world application that would help a farmer recognizing a specific weed among thousands of potential species. Our claim is actually that more diverse views of the plants, fitting the diversity of much richer floras, should actually be considered to achieve a real ecological impact. The use of leaves alone has actually many practical and botanical limitations. Leaves are typically not visible all over the year for a large fraction of plant species and for many others, they are intrinsically not enough informative or very difficult to capture (needles of pines, grass, huge leaves of banana trees...). Using flowers alone, as suggested in [17, 1], is not a better choice because these limitations are even more accentuated. The mobile application presented in this demo is the first image-based identification tool dealing with multiple organs of the plants (leaf, flower, fruit and bark) in their natural environment. It allows querying the system at any period of the year and it benefits from the complementarities of the different views of the plant to improve identification performances.

Pl@ntNet mobile application is built on top of an end-to-end collaborative information system allowing the training data to be continuously enriched and revisited [15]. Complementary to world-wide data integration efforts such as *Encyclopedia Of Life*[3] or crowd-sourced approaches such as ImageNet [5] or NEC flowers dataset [1], our proposal is to rely on thematic social networks to solve data validation issues and produce accurate knowledge. Epistemic communities actually have the advantage to connect experts, enlightened enthusiasts, amateurs and novice around the same topic so that all of them can play complementary roles in a real-world ecological surveillance workflow. Experts can typically animate projects, define observation protocols and teach, amateurs can collaboratively validate data by interacting together according to their level of expertise, novice can provide massive sets of observations.

More concretely, our workflow relies on Tela Botanica[4], a French-speaking network linking 20K members in more than 70 countries (41% *novice* in botany, 30% *with a good practice* and 7% *experts*). Raw image feeds, collected through web crawls, personal collections or through the mobile fronts themselves, are integrated in a collaborative tool so that they can be enriched and validated by the network. This

[1] `https://itunes.apple.com/en/app/plantnet/id600547573`

[2] http://leafsnap.com/

[3] http://www.eol.org/

[4] http://www.tela-botanica.org/

ICMR 2014

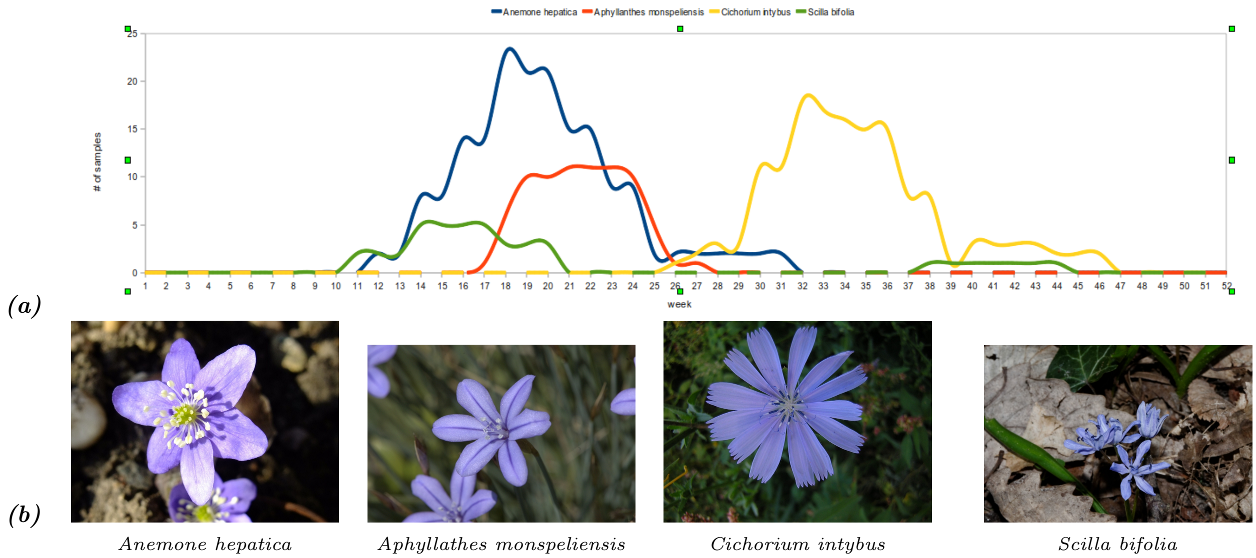

Figure 1: Flowering periods for sample images with similar color and/or texture

typically allows managing contradictory determinations that result from the intrinsic complexity of identifying a plant (classical dichotomous keys used by botanists require answering dozens or even hundreds of complex questions related to thousands of often ambiguous morphological attributes). To boost integration, contents with missing organ tags are automatically annotated and can be corrected afterwards. Based on all the produced metadata, only relevant contents are finally inserted daily in the visual index.

2. THE VISUAL SEARCH ENGINE

A detailed description of Pl@ntNet visual search engine is available in [15]. Its core algorithm is built from the works of Joly et al. [13, 12, 14] on large-scale matching. It was actually shown in [8, 9] that such object retrieval methods do work surprisingly well on plants. Concretely, image retrieval is achieved through the following steps: (i) Local features extraction (ri-LBP [18]; SURF [3], a 20-dim. Fourier histogram [6]; an 8-dim. Edge Orientation Histogram, a 27-bin weighted RGB histogram, and a 30-bin HSV histogram; all extracted around multi-resolution color Harris points) (ii) Hamming embedding and indexing with RMMH [14] (iii) approximate k-nn search with AMP-LSH [12] (iv) image-level scoring by counting the number of matched features weighted by their distances to query features. A multi-stage late fusion strategy is then used to efficiently combine the different features and the multiple query images (see section 3.4).

3. THE NEW FEATURES

3.1 Android port

The first iPhone version of the application was released in March 2013 and was enriched with the following new features as well as with many other minor improvements concerning the interactive GUI, the correction of bugs, etc. The Android version, introduced in this paper, is currently being evaluated by 314 beta testers and will be freely available on Google Store before Spring 2014. Both versions interact with the same remote visual search engine (in a client/server manner) and the same data store. Both of them also share the same core functionalities: (i) an image feeds reader to explore the last contributions of the community (ii) a taxonomic browser with full text search options (iii) a user profile and personal contents management screen and (iv) the image-based identification tool itself. This one first asks the user to take a picture and then let him chose among 4 icons representing a flower, a leaf, a fruit and a bark. Up to five pictures of the same plant can be acquired in this way and the complete set of pictures can finally be submitted as a query plant to the remote visual search engine. Retrieved species with confidence scores and matched images are finally returned to the device and displayed on the the result screen by decreasing confidence. Selecting one of the retrieved species opens a detailed view screen with all matched pictures (classified by organ galleries) allowing a first stage of refinement in the determination process. A second stage of refinement can be achieved by accessing either eFlore fact sheets (the most complete db on France flora) or wikipedia mobile pages. If the user believes having found the right species, he can finally contribute by sending his observation with pictures, date, gps and author's name (under Creative Commons license). The observation will instantaneously join a collaborative web tool[5] allowing the community to revise and rate it before it can be inserted in the reference set.

3.2 More data

Thanks to new contributions and integration efforts, the indexed image collection now contains 84,525 images covering 3,931 species (to be compared to 22,574 pictures representing 957 species in [7]). To our knowledge, there is no other identification tool covering a so large number of species. As a comparison, LeafSnap [16] application contains

[5] `http://www.tela-botanica.org/appli:identiplante`

only 184 tree species in the Northeastern United States. Encyclopedic initiatives like EOL[6] and MorphBank[7] cover much more species but must often with only very few pictures per organ and per species and no practical identification functionalities. On the other side, crowdsourced datasets such as ImageNet [5] contain thousands of pictures for a lot of species, but still without organ tags and with a very high level of noise for a botanical usage. More generally, most existing data suffer from a *long tail* distribution, i.e. with very few species well represented and most species with very few images.

3.3 Filtering by flowering period

For many species, flower organs are present for just a quite short period of time, and each species has its own flowering periods. As the metadata of the collected Pl@ntNet images contains the date when the photograph was taken, we can apply a post-processing treatment applied to the list of the purely visual results. The idea is that species that do not contain any observation in the dataset matching the period of the query observation are pruned. More precisely, the flowering period histogram for a species is constructed by week, with ± 3 additional weeks to account for geographical and year-to-year differences. Given a training image of class C^k, taken in week w, histogram bins $H^{C^k}(h), h = w-3, ..., w+3$ are incremented. For a query image Q taken in week w^Q, an histogram H^Q is constructed in the same manner, and finally a species C^k is retained if $\exists w | H^Q(w) > 0 \wedge H^{C^k}(w) > 0$.

Figure 1 shows the flowering periods for four species that have similar color and texture. *Cichorium intybus* flowers appear cleary later than the other species over a year. Thus, any query images in this period will exclude the three other species, even if the visual content are very similar to these species.

3.4 Multi-feature and multi-image fusion

We extended the two-stage late fusion module introduced in [15] to allow it fuse the different visual features rather than concatenating them in a early fusion manner before applying RMMH. We still index the different view types separately but we now also build separate visual indexes for each feature type. Consequently, the new late fusion algorithm includes three steps: a multi-features level, a multi-image level and a multi-view level. The new feature-level stage first converts the results lists of each feature into species distributions and then merge them according to their relative discrimination. The detailed description of the full fusion module can be found in our ImageCLEF 2013 working note [2].

3.5 cross-languages functionalities

Whereas the latin taxonomic name of a plant is language independent, its common name(s) vary a lot from a country to another. On the other side, this was the most requested missing functionality by the users of the application after its first release. We therefore achieved a consistent data integration effort to find existing thesauri in 7 European languages and align them with our taxonomic reference (the French official one). The language of the application is now automatically switched according to the device settings (with english used as default if the device language does not matched any of the 7 we integrated).

[6] `eol.org`

[7] `http://www.morphbank.net/`

4. NEW EXPERIMENTS

4.1 Participation to ImageCLEF 2013

Our system was evaluated within ImageCLEF 2013 international evaluation campaign that includes an image-based plant identification task [10]. The task consisted in providing a ranked list of species for 5,092 unknown plant images and a training set of 20,985 images covering about 250 species of plants living in France area. Each picture was associated with a set of metadata including the view type (Entire plant, Flower, Fruit, Leaf and Stem), the date and the location of the observation, a unique observation identifier and the author name. Figure 2 presents the overall results achieved by the 12 participating groups who submitted a total of 33 runs. Pl@ntNet system with all new features (*Inria PlantNet Run2*), obtained the second best result with an average weighted classification score of 0.385 average (see [10] for more details on the evaluation protocol and the metric). Our second best run (*Inria PlantNet Run1*) correspond to a purely visual run and allows to measure the benefit of using the metadata in the identification process by comparing it to *Run2*. Table 1 presents the detailed results for each view type (only for the top-5 runs of the task). It shows how the use of the flowering period boosts the performances for the flower category (between our textitRun1 and *Run2*).

Run	Entire	Flower	Fruit	Leaf	Stem
NlabUTokyo Run3	**0.297**	0.472	**0.311**	**0.275**	0.253
Inria PlantNet Run2	0.274	**0.494**	0.260	0.272	0.240
NlabUTokyo Run2	0.273	0.484	0.259	0.273	**0.285**
Inria PlantNet Run1	0.254	0.437	0.249	0.240	0.211
NlabUTokyo Run1	0.236	0.423	0.209	0.269	0.276

Table 1: ImageCLEF 2013 - Results of the plant identification task (natural background)

4.2 Real-world experiments and user feedbacks

As the iPhone version of our application was released in March 2013, we collected many user statistics and feedbacks that helped us evaluating the application and improving it. At the time of writing, 75 525 people from did download the application from 142 countries. The weekly number of active users ranges from about 1000 to 5000 users depending on the period of the year. The daily number of sessions depends on the weather and the day of the week (with periodic peaks on sundays). The number of crashes decreased from 600 for the first release (1.0.1) to 96 for the last one (1.0.5) for a total of 407,999 sessions. We collected 123 customer reviews from the French iTune portal of the application (and 10 from the english one) with an average note of 3 stars. 37% of users gave a 5 or 4 stars note and very positive feedbacks ("excellent", "genial", "super" etc.). On the other side, 47% of users gave the lowest note (i.e 1 star). Their comments suggest that they tried one or two queries and gave up because the application did not return the write answer

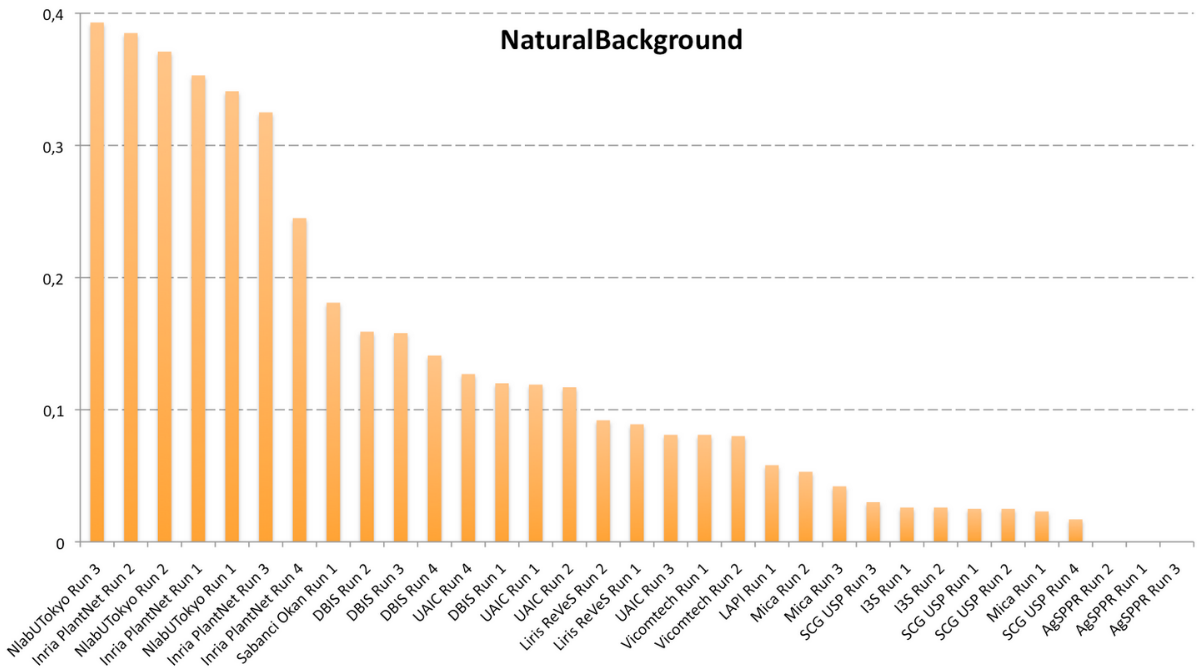

Figure 2: ImageCLEF 2013 - Results of the plant identification task (natural background)

("does not work", "did return random results for a picture of banana", "not mature enough", etc.). We know from the logs of the application that a large part of these unsuccessful identifications are related to out-of-scope queries (e.g. ornamental or horticultural plants whereas the application is restricted to wide plants so far) or query pictures of poor quality (blurred pictures erasing informative details, very small object of interest with high clutter, etc.). Other failures are more related to the intrinsic performances of the system (in particular species with too few images in the training set).

5. REFERENCES

[1] A. Angelova, S. Zhu, Y. Lin, J. Wong, and C. Shpecht. Development and deployment of a large-scale flower recognition mobile app, December 2012.

[2] V. Bakic, S. Mouine, S. Ouertani-Litayem, A. Verroust-Blondet, I. Yahiaoui, H. Goëau, and A. Joly. Inria's participation at ImageCLEF 2013 Plant Identification Task, Sept. 2013. Pl@ntNet (Agropolis fondation).

[3] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Surf: Speeded up robust features, 2008.

[4] G. Cerutti, L. Tougne, A. Vacavant, and D. Coquin. A parametric active polygon for leaf segmentation and shape estimation, 2011.

[5] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database, 2009.

[6] M. Ferecatu. Image retrieval with active relevance feedback using both visual and keyword-based descriptors. PhD thesis, Univ. Versailles St-Quentin, 2005.

[7] H. Goëau, P. Bonnet, A. Joly, V. Bakić, J. Barbe, I. Yahiaoui, S. Selmi, J. Carré, D. Barthélémy, N. Boujemaa, J.-F. Molino, G. Duché, and A. Péronnet. Pl@ntnet mobile app, 2013.

[8] H. Goëau, P. Bonnet, A. Joly, N. Boujemaa, D. Barthélémy, J.-F. Molino, P. Birnbaum, E. Mouysset, and M. Picard. The ImageCLEF 2011 plant images classification task, 2011.

[9] H. Goëau, P. Bonnet, B. Julien, V. Bakic, A. Joly, and J.-F. Molino. Multi-Organ Plant Identification, 2012.

[10] H. Goëau, A. Joly, P. Bonnet, V. Bakic, D. Barthélémy, N. Boujemaa, and J.-F. Molino. The imageclef plant identification task 2013, 2013.

[11] C. Grozea. Brainsignals Submission to Plant Identification Task at ImageCLEF 2012, 2012.

[12] A. Joly and O. Buisson. A posteriori multi-probe locality sensitive hashing, 2008.

[13] A. Joly and O. Buisson. Logo retrieval with a contrario visual query expansion, 2009.

[14] A. Joly and O. Buisson. Random maximum margin hashing, 2011.

[15] A. Joly, H. Goeau, P. Bonnet, V. Bakić, J. Barbe, S. Selmi, I. Yahiaoui, J. Carré, E. Mouysset, J.-F. Molino, N. Boujemaa, and D. Barthélémy. Interactive plant identification based on social image data. Ecological Informatics, (0):–, 2013.

[16] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. B. Soares. Leafsnap: A computer vision system for automatic plant species identification, 2012.

[17] M.-E. Nilsback and A. Zisserman. Automated flower classification over a large number of classes, 2008.

[18] T. Ojala, M. Pietikäinen, and T. Mäenpää. Gray scale and rotation invariant texture classification with local binary patterns. In Computer Vision - ECCV 2000, volume 1842 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2000.

[19] C. Spampinato, V. Mezaris, and J. van Ossenbruggen. Multimedia analysis for ecological data, 2012.